Journal of Biomolecular NMR, **23:** 157–158, 2002.
KLUWER/ESCOM

Letter to the Editor: Assignment of the ^{1}H, ^{13}C, and ^{15}N resonances of the 22,5 kDa CBM28 module of the cellulase Cel5I of *Clostridium cellulolyticum*

Amor Mosbah[a], Chantal Tardif[b,d], Olivier Bornet[c], Odile Valette[b], Bernard Henrissat[a] & Hervé Darbon[a,*]

[a]*Architecture et Fonction des Macromolécules Biologiques CNRS UMR 6098, France;* [b]*Bioénergétique et Ingénierie des Protéines, CNRS UPR 9036, France;* [c]*Institut de Biologie Structurale et Microbiologie, 31 Chemin Joseph-Aiguier, 13402 Marseille Cedex 20, France;* [d]*Université de Provence, France*

Received 5 February 2002; Accepted 20 March 2002

Key words: CBM28, Cel5I, chemical shifts, *Clostridium cellulolyticum*, heteronuclear NMR

Biological context

Cellulolytic microorganisms produce highly efficient enzyme systems capable of degrading cellulose, the most abundant but also probably the most recalcitrant polysaccharides on Earth. Some microorganisms, notably the aerobes, produce soluble modular enzymes composed of a catalytic module carrying one or several non-catalytic modules, often functioning as carbohydrate-binding modules (CBMs), on a single polypeptide chain. In contrast, the cellulase systems of anaerobic bacteria and fungi are multi-component protein assemblies (several mega-Daltons molecular weight) named 'cellulosomes' (Lamed et al., 1983; Shoham et al., 1999). The enzymes are incorporated into the multisubunit complex via a cellulosome integrating protein named scaffoldin (Tokatlidis et al., 1993; Bayer et al., 1994). Specialized complementary modules, the multiple copies of the cohesins on the scaffoldin, and the single dockerin module on each enzyme, are the elements responsible for cellulosome assembly (Salamitou et al., 1994; Pagès et al., 1996).

In addition to the dockerins, cellulosomal enzymes sometimes contain CBMs. The CBMs have been classified in distinct families based on amino acid sequence similarities (Tomme et al., 1995). Currently there are 29 known families of CBMs (these are conveniently accessed via a www server at http://afmb.cnrs-mrs.fr/~cazy/CAZY/CBM.html). The cellulase Cel5I of *Clostridium cellulolyticum* is a highly modular protein with a catalytic domain belonging to glycoside hydrolase family GH5, carrying a CBM from family CBM17, a CBM from family CBM28 and three S-layer homology domains (GenBank access code: AY077754).

All occurrences of the CBM28 modules are within modular proteins involved directly or indirectly in cellulose degradation (Boraston et al., 2002). Currently there is no 3-D structure available for any CBM of family CBM28. Here we report the backbone and side chain $^{13}C^{\beta}$ resonance assignment of the CBM28 module (198 amino acid residues) of the cellulase Cel5I of *Clostridium cellulolyticum*.

Methods and results

Sample preparation

The CBM28 encoding sequence was amplified by PCR from pETI1, the expression vector coding for the Cel5I entire form of *Clostridium cellulolyticum* (unpublished data), the forward 5′-*CCCTATACATATG*ACTACCGTAGAAGCACCTGT TG-3′ (*Nde*I site underlined, ATG codon in bold), and the reverse 5′-*GTGCTCGAG*GCTGTCATTTTCGA ACCTGAC-3′ (*Xho*I site underlined) primers having partial homology with the CBM28 encoding DNA (upper case). The *Nde*I-*Xho*I synthesised fragment was cloned into *Nde*I-*Xho*I linearised pET22b(+), thereby generating pETCBM28. The His-tagged CBM28 module was produced by the recombinant *Escherichia coli* BL21(DE3) [pETCBM28] strain, which was grown at 37 °C to $OD_{600} = 2$ in Luria–Bertani medium, supplemented with 1.2% glycerol (w/v) and ampicillin (200 μg ml^{-1}), and then incubated at 18 °C overnight with isopropyl thio-β-D-galactoside (IPTG) 50 μM. The purification of CBM28 was performed

*To whom correspondence should be addressed. E-mail: herve@afmb.cnrs-mrs.fr

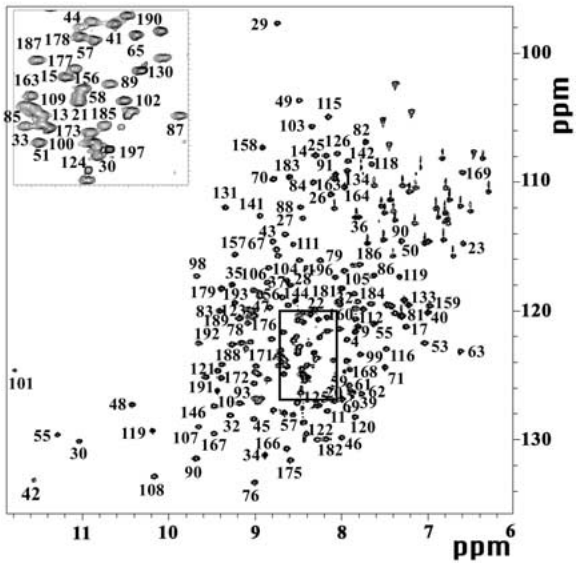

Figure 1. 500 MHz ^{1}H-^{15}N HSQC spectrum of ^{15}N-CBM28 with full assignment, showing the well-dispersed amide proton resonances. Side chain amino groups are not labelled.

on Ni-NTA resin, as previously described (Reverbel-Leroy et al., 1997) except that the protein was eluted with a linear gradient from 5 to 100 mM imidazole in 30 mM Tris/HCl, pH. The ^{15}N-labeled and the ^{13}C/^{15}N-labeled proteins were produced in mineral medium containing ^{15}N NH_4Cl 0.8 g l^{-1} + unlabeled glycerol 15 g l^{-1}, and ^{15}N NH_4Cl 0.8 g l^{-1} + ^{13}C glycerol 2 g l^{-1}, respectively. Cells were grown at 37 °C to OD_{600} = 1. Induction was, in this case, performed overnight at 25 °C with 100 μM of IPTG.

NMR spectroscopy

Homonuclear NMR experiments were performed with 4.5 mM of unlabelled protein sample at 300 K, in 10 mM sodium deuterated acetate buffer at pH 5.0 containing 10% D_2O (v/v). All NMR spectra were recorded on a 500-MHz DRX Bruker.

Heteronuclear NMR experiments were performed on a 3.5 mM uniformly ^{15}N-labelled protein in 10 mM sodium deuterated acetate buffer at pH 5.0 containing 10% D_2O (v/v). HSQC, 2D HSQC-NOESY and 2D HSQC-TOCSY, 3D NOESY-HSQC and 3D TOCSY-HSQC, and 3D HSQC-NOESY-HSQC spectra were acquired using the Fast HSQC scheme (Mori et al., 1995). ^{15}N decoupling during acquisition was achieved using the GARP sequence. The 2D HSQC-NOESY, 3D NOESY-HSQC and 3D HSQC-NOESY-HSQC were performed with a mixing time of 100 ms. A 80 ms DIPSI-2 mixing sequence was used in the 2D HSQC-TOCSY and 3D TOCSY-HSCQ experiments. 3D HNCA, HN(CO)CA, HNCO, HN(CA)CO, CBCANH, CBCA(CO)NH, HBHA(CBCACO) NH, HBHA(CBCA)NH and H(C)CH-TOCSY spectra were acquired with 2.1 mM of uniformly ^{15}N-^{13}C-labelled protein.

Extent of assignment and data deposition

With the combined information of all heteronuclear experiments we were able to assign 90% of the backbone ^{15}N and NH. The unassigned residues are mostly located at the proximity of prolines (5–7; 59–65; 71–73; 93–97; 125–130; 112–115) and the stretch 147–150. Prolines ^{13}Cα and ^{13}Cβ resonance were determined from the CBCACONH. The ^{15}N-HSQC spectrum of CBM28 at 300 K is shown in Figure 1. A detailed examination of the ^{13}Cα resonance frequencies and the existence of long range nOes indicate that CBM28 is mainly organized in β-strands.

The ^{1}H, ^{15}N and ^{13}C chemical shifts for CBM28 at 300 K have been deposited in the Bio-MagResBank (http//:www.bmrb.wisc.edu) under BMRB accession number 5263.

Acknowledgement

One of us (Amor Mosbah) is a recipient of a Ph.D. fellowship from CNRS.

References

Bayer, E. A., Morag, E. and Lamed, R. (1994) *Trends Biotechnol.*, **12**, 378–386.

Boraston, A.B., Ghaffari, M., Warren, R.A. and Kilburn, D.G. (2002) *Biochem. J.*, **361**, 35–40.

Lamed, R., Setter, E., Kenig, R. and Bayer, E.A. (1983) *Biotechnol. Bioeng. Symp.*, **13**, 163–181.

Mori, S., Abeygunawardana, C., Johnson, M. and Van Zijl, P.C.M. (1995) *J. Magn. Reson.*, **B108**, 94–98.

Pagès, S., Belaich, A., Tardif, C., Reverbel-Leroy, C., Gaudin, C. and Belaich, J.P. (1996) *J. Bacteriol.*, **178**, 2279–2286.

Reverbel-Leroy, C., Pagès S., Bélaich, A., Bélaich, J.P. and Tardif, C. (1997) *J. Bacteriol.*, **179**, 46–53.

Salamitou, S., Raynaud, O., Lemaire, M., Coughlan, M., Béguin, P. and Aubert, J.P. (1994) *J. Bacteriol.*, **176**, 2822–2827.

Shoham, Y., Lamed, R. and Bayer, E.A. (1999) *Trends Microbiol.*, **7**, 275–281.

Sunna, A., Gibbs, M.D. and Bergquist, P.L. (2001) *Biochem. J.*, **356**, 791–798.

Tokatlidis, K., Dhurjati, P. and Béguin, P. (1993) *Protein Eng.*, **6**, 947–952.

Tomme, P., Warren, R.A., Miller, Jr., R.C., Kilburn, D.G. and Gilkes, N.R. (1995) In *Enzymatic Degradation of Insoluble Carbohydrates*, Saddler, J.N. and Penner, M.H. (Eds.), American Chemical Society, Washington, DC.